

VU Research Portal

Stochastic aspects of familial searching

Kruijver, M.V.

2016

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Kruijver, M. V. (2016). *Stochastic aspects of familial searching*. [, Vrije Universiteit Amsterdam].

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

Contents

1	Introduction	9
1.1	Forensic genetics	9
1.1.1	DNA profiles	10
1.1.2	Genotype probabilities	11
1.1.3	Mendelian inheritance	12
1.1.4	Paternity testing	13
1.1.5	Evidence interpretation using likelihood ratios	14
1.1.6	Pedigree likelihoods	15
1.1.7	Likelihood ratio distributions	17
1.2	Outline of the thesis	18
2	Optimal strategies for familial searching	21
2.1	Introduction	21
2.2	Four search strategies	22
2.3	Theoretical aspects	24
2.4	Empirical performance of search strategies	28
2.5	Applying a search strategy	32
2.5.1	Fixing the PoD	33
2.5.2	Fixing the posterior probability of a relative	35
2.5.3	Conclusion	36
2.6	Heterogeneous databases	37
2.7	Composite searching	39
2.7.1	Bycatch of simple searches	39
2.7.2	Three composite search strategies	40
2.7.3	Implementation	41
2.7.4	Efficiency and misspecification of the prior	43
2.8	Conclusion	46
2.9	Appendix: distribution of KI	47
3	Computations and importance sampling	49
3.1	Introduction	49
3.2	Problem description	51
3.2.1	Example (likelihood ratio for pairwise kinship with a fixed person)	52
3.2.2	Example (likelihood ratio for pairwise kinship for two persons)	53

3.3	Approaches from the literature	54
3.3.1	Simulation	54
3.3.2	Normal approximation	55
3.3.3	Corradi and Ricciardi's approach	55
3.3.4	Dørum et al.'s algorithm for tail probabilities	56
3.4	Efficient simulation with importance sampling	57
3.5	A new efficient algorithm for exact computation	58
3.6	A new approach using approximation	60
3.7	Applications	61
3.7.1	Familial searching: case pre-assessment	61
3.7.2	Familial searching: a stringent strategy	63
3.7.3	Power of kinship identification systems: full sibs versus half sibs	63
3.7.4	A two-person mixture	64
3.8	Conclusion	66
4	Characterizing genetic structure	67
4.1	Introduction	67
4.2	Product rule and Balding-Nichols	69
4.3	Latent variable model	71
4.3.1	Introduction	71
4.3.2	Model specification and estimation	71
4.3.3	Fitting the model to the Dutch database	74
4.4	Application 1: predicting false leads	77
4.4.1	Familial searching	77
4.4.2	Searches for a contributor to a mixture	83
4.4.3	Conclusions	84
4.5	Application 2: partial matching	86
4.5.1	Matching and partially matching profiles	86
4.5.2	Examining the Dutch database	88
4.6	Application 3: validating BN match probabilities	91
4.6.1	Match probabilities	91
4.6.2	Empirical study	92
4.7	Discussion	93
4.8	Appendix: Dutch database	94
5	P-values of the likelihood ratio	101
5.1	Introduction	101
5.2	The likelihood ratio approach	102
5.3	P-values of the likelihood ratio in mixture cases	103
5.3.1	Example: p-value suggests stronger evidence than likelihood ratio	104
5.3.2	Example: p-values may remain small for exclusionary likelihood ratio	105
5.4	P-values are ambiguous	108
5.4.1	Example (full sibling test)	108
5.5	Further discussion and conclusion	110

6	Familial searching with more markers	113
6.1	Introduction	113
6.2	Likelihood ratio distributions	114
6.3	Top- k probabilities	117
6.3.1	A database of one million profiles	118
6.3.2	Larger databases	118
6.4	Probabilities of exceeding LR thresholds	119
6.4.1	$LR > 1,000$	120
6.4.2	Higher thresholds	121
6.5	Half-siblings and further relatives	122
6.5.1	More loci	123
6.6	Conclusion	124
6.7	Appendix: methodological details	125
	Bibliography	127